**How to run Batch HMM classification**

The HMM is vbSPT and I always use a 2-state version: <http://vbspt.sourceforge.net/>

Everything is in this folder: /Users/anderssejrhansen/Dropbox/Shared\_with\_Assaf/AndersScripts/HMM\_classification

Please open the script “Batch\_vbSPT\_classify.m” and press run. Everything should run on the example data and expect it to take around 5 min to finish.

The script reads in the frame rate from the file name. E.g. “20171023\_sim\_mESC\_C45\_mRad21-Halo\_G1\_133Hz\_pooled\_QC\_CD2.mat”. So please make sure you name your simulations either 223Hz, 133Hz or 74Hz.

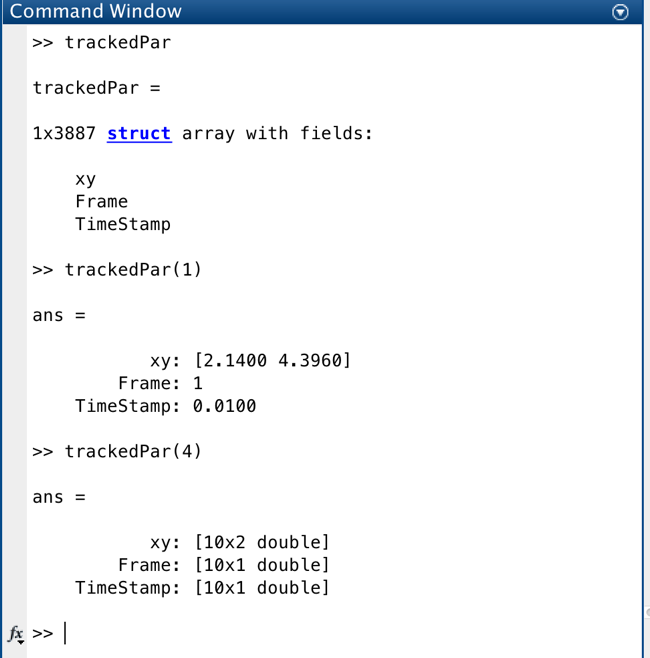
Here is what the batch script does:

* It first reads in SPT data in my own preferred format. In “TestData”, I put 2 example SPT files (.mat format). They should contain a structured array called “trackedPar\_QC”. Please use this format:

The MAT-files contain a [structure array](https://www.mathworks.com/help/matlab/ref/struct.html) named “trackedPar\_QC”. trackedPar\_QC contains three variables:

* trackedPar\_QC.xy: “xy” is a matrix with 2 columns and a number of rows corresponding to the number of localizations in that trajectory. The first column is the x-coordinate and the second column is the y-coordinate. The units are micrometers (μm).
* trackedPar.Frame\_QC: “Frame” is a column vector where each element is the frame where the particle was localized.
* trackedPar.TimeStamp\_QC: “TimeStamp” is a column vector where each element is the timepoint where the particle was localized.

Each element in the structure array “trackedPar\_QC” correspond to a different trajectory as illustrated below:



* Once it has read in the data, the data gets converted to a cell array format preferred by vbSPT and saved in the folder “QC\_data\_reformatted”
* Now it’s ready to do the HMM classification. vbSPT annoyingly reads a file called “vbSPT\_RunInputFileBatch.m”. So I custom wrote a script to edit these files and copy them to the folder “FolderWithRunInputFiles”. As long as you keep the format consistent, you should not need to deal with this.
* It will then perform HMM-classification (about 2-3 min per file) and save the HMM-classificed data in “QC\_HMM\_classified\_data”.
* The final MAT-file contains: “CellTracks”, a cell-array of x-y coordinates and “CellTracksViterbiClass”, the HMM-classification with 1 for bound and 2 for free.